



# **Collecting Data Electronically**

Keeping QA Staff Bored

November 14, 2018

COMMITTED TO SAFE, CLEAN, ENJOYABLE CREEKS AND RIVERS

# Collecting Data Electronically

- Recording initial observations in an electronic format, including via:
  - Phone/Tablet App
  - Spreadsheet on a PC
  - Specialized program or Instrument
- Can include collection through final storage
- In-house, out-of-the-box, or a customized commercial program



# Why?

## Paper is easy, but...

- Waterproof paper and pens get pricey
- Mistakes get messy, further complicating validation and corrections
- Long term storage

SAN ANTONIO RIVER AUTHORITY  
Field Data Sheet

Check for Review by Field Staff  
(Initial/Date) *CE 9/27/20*

Sample No.(s): \_\_\_\_\_ Matrix: ☒ NFW ☐ QC ☐ Sediment

Station Id: **TCEQ-21547** Station Location: **SAT @ VPW St. SARC View Blvd. 8/12/18**

Program Code: ☒ CRP ☐ SARA\_SM ☐ TXDOT\_Storm ☐ PC ☐ SAR\_Loop ☐ NPS\_Project ☐ Storm

☐ Special Request (Specify Requestor in Comments) ☐ (Please Specify)

Sample Type: ☒ RRT ☒ BS ☐ BF ☐ AS ☐ BE ☐ Pollution Complaint ☐ SR ☐ SS (Specify in Comments) ☐

Monitoring Type 2: **RT** Collection Method: ☐ Grab Instrument #: \_\_\_\_\_

Submitting Entity: SA *9-12-18* Collecting Entity: *AS 9-12-18* 13.12 *CE*

Collection Date: *9-12-18* Collection Time: *01:12 9-12-18* End Depth: *0.7*

Collector(s) Signature(s): *[Signature]*

# of Containers/Container Type	Type of Field Preservation	Requested Analysis	pH <2 (Y or N)?
1 GC - Gallon Cubitainer	<input checked="" type="checkbox"/> Ice <input type="checkbox"/> H <sub>2</sub> SO <sub>4</sub> <input type="checkbox"/> HNO <sub>3</sub> <input type="checkbox"/> Filtered	NO <sub>2</sub> N, NO <sub>3</sub> N, TSS, CL, <i>SO<sub>4</sub></i>	
1 QC - Quart Cubitainer	<input checked="" type="checkbox"/> Ice <input type="checkbox"/> H <sub>2</sub> SO <sub>4</sub> <input type="checkbox"/> HNO <sub>3</sub> <input type="checkbox"/> Filtered	TPO <sub>4</sub> , NH <sub>4</sub> , TKN	
1 LW - Large Whirlpak	<input checked="" type="checkbox"/> Ice <input type="checkbox"/> H <sub>2</sub> SO <sub>4</sub> <input type="checkbox"/> HNO <sub>3</sub> <input type="checkbox"/> Filtered	E. COLI	
1 PB - Plastic Bottle	<input checked="" type="checkbox"/> Ice <input type="checkbox"/> H <sub>2</sub> SO <sub>4</sub> <input type="checkbox"/> HNO <sub>3</sub> <input type="checkbox"/> Filtered	METALS, CALC. HARDNESS	
Other (Specify) _____	<input type="checkbox"/> Ice <input type="checkbox"/> H <sub>2</sub> SO <sub>4</sub> <input type="checkbox"/> HNO <sub>3</sub> <input type="checkbox"/> Filtered		
Other (Specify) _____	<input type="checkbox"/> Ice <input type="checkbox"/> H <sub>2</sub> SO <sub>4</sub> <input type="checkbox"/> HNO <sub>3</sub> <input type="checkbox"/> Filtered		

☒ Field Parameters ☐ Flow ☐ H<sub>2</sub>SO<sub>4</sub> Reagent ID: R \_\_\_\_\_

☐ Observed/corrected temperature (ID: C01-096) ☐ pH Paper: R01-003 Sample Temperature (°C): *1*

Associated Required Data: ☐ Nekton ☐ Benthic ☐ Habitat ☐ 24hr DO ☐ Metals ☐ Other (Specify in Comments)

**FIELD PARAMETERS**

Meter: *600 XLM* ☒ 600V2-2M ☒ EXO1 ☐ EXO2

Parameter	Code	Value
Dissolved Oxygen	00300	<i>7.17 8/12/18 mg/L</i>
Temperature	00010	<i>23.1 °C</i>
pH	00400	<i>8.1 7.8 S.U.</i>
CL <sub>2</sub>	50060	<i>0.8 0.2 mg/L</i>
*Conductivity (temperature compensated value to 25 °C)	00094	<i>298.1 µS/cm</i>
Secchi Depth	00078	<i>2.2 m</i>
Days Since Last Precipitation Event	72053	<i>2.1 days</i>

**FIELD OBSERVATIONS**

Depth Sensor reset to 0.00 (prior to collecting first sample of the day): ☒

Parameter	Code	Observations
Flow Severity	01351	<input checked="" type="checkbox"/> 1 - No Flow <input type="checkbox"/> 3 - Normal <input type="checkbox"/> 5 - High <input type="checkbox"/> 2 - Low <input type="checkbox"/> 4 - Flood <input checked="" type="checkbox"/> 6 - <i>12-18</i>
Water Color	89969	<input type="checkbox"/> 1 - Brown <input type="checkbox"/> 3 - Green <input checked="" type="checkbox"/> 5 - Clear <input type="checkbox"/> 2 - Reddish <input type="checkbox"/> 4 - Black <input type="checkbox"/> 6 - Other (Specify in Comments)
Water Odor	89971	<input type="checkbox"/> 1 - Sewage <input type="checkbox"/> 3 - H <sub>2</sub> S <input type="checkbox"/> 5 - Fishy <input type="checkbox"/> 7 - Other (Specify in Comments) <input type="checkbox"/> 2 - Oily / Chemical <input type="checkbox"/> 4 - Musky <input checked="" type="checkbox"/> 6 - None
Present Weather	89966	<input type="checkbox"/> 1 - Clear <input checked="" type="checkbox"/> 3 - Cloudy <input type="checkbox"/> 2 - Partly Cloudy <input type="checkbox"/> 4 - Rain
Contact Recreation Observations	SA699	<input type="checkbox"/> 1 - Primary Observed <input type="checkbox"/> 3 - Non-contact Observed <input type="checkbox"/> 6 - Non-Contact evidence <input type="checkbox"/> 2 - Secondary Observed <input type="checkbox"/> 4 - Primary Evidence <input checked="" type="checkbox"/> 7 - No Evidence <input type="checkbox"/> 5 - Secondary Evidence
# of People Observed	89978	Evidence of Primary Contact Recreation <i>0</i> 89979 <input type="checkbox"/> Observed (1) <input checked="" type="checkbox"/> Not observed (0)

Field Comments: *Site noted pooled 8/12/18*

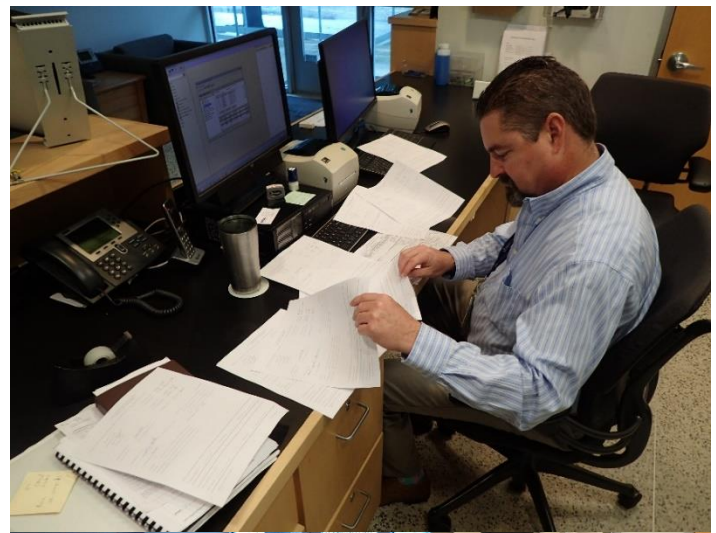
F007\_Rev20 Issued by SARA QA: JH; Effective: 07/30/2018 Page 1 of 2



COMMITTED TO SAFE, CLEAN, ENJOYABLE CREEKS AND RIVERS

# Why?

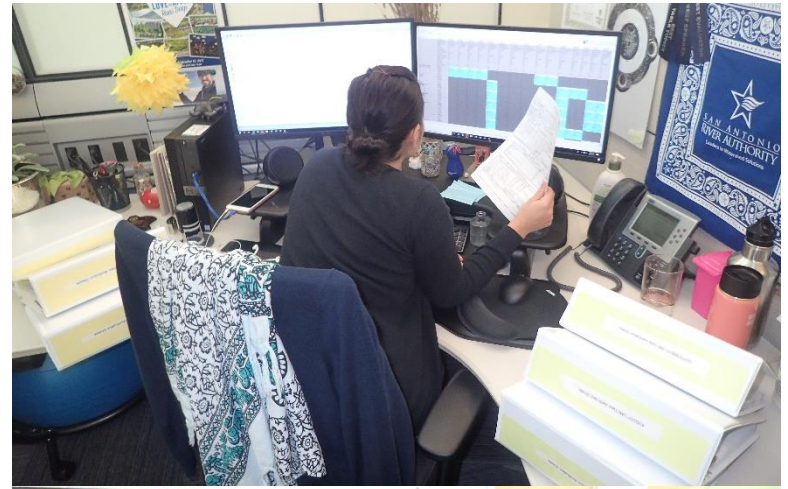
- Much faster overall
- Simpler COC and easier log-in
  - Fewer errors in preparing COCs and at log-in.
- Automatic calculations
  - Calculations completed in background as data is collected; no more calculation errors or method differences





# Why?

- Live checks for user-errors and completeness
- Export file for analysis or database import
  - “Hands off” data entry (entry is processed by LIMS via CSV)
  - Validation still necessary, but much quicker
  - Entry accuracy virtually 100%



# Why?

- Other Benefits

- Less paper

- RT samples = 770 sheets; electronic = 124 sheets; **84% reduction**
    - Biologicals = 348 sheets; electronic = 84 sheets; **76% reduction**

- Time savings (prep through validation)

- **10-15 minutes** per RT sample
    - **3-4 hours** for a single biological event
    - **Total of 210-300 man-hours saved per year**





# Robustness needed

- Depends on a number of factors
  - Staff abilities and habits
  - Errors reduced, minimalized or *eliminated*?
  - Agency support
    - QA, budgeting for equipment, user acceptance
  - Acquisition source?
  - Maintenance plans?



# Who?

- Acquisition
  - **In-house:** understand full process and be capable; customized to YOU; may be dependent on one person
  - **Out-of-the-box:** quick and easy, may have to modify your methods to fit; customized to ALL
  - **Custom commercial:** \$\$\$; customized to your needs with limitations; locked-in
- Maintenance plans
  - Who? Price? How often?





# Cautions

	Error	Bug
When	Original Entry	Data handling
How	User error	Bug in application
Caught	Immediately, post-processing, or during QA	QA; experienced staff
Example	Dissolved oxygen = 81 instead of 8.1; missed datapoint	Calculation error (e.g., “Not Recorded” processed as 0)



# Cautions

- User-error ~~proofing~~ resistant
  - Address ALL possible user-errors (ignore, prevent, or handle)
  - Which errors are worth the trouble? (e.g. DO = orange)
  - When necessary, account for the occurrence (prevent or handle appropriately)
  - *Be careful not to inadvertently exclude valid data*



# Cautions, cont.

- Bug Prevention

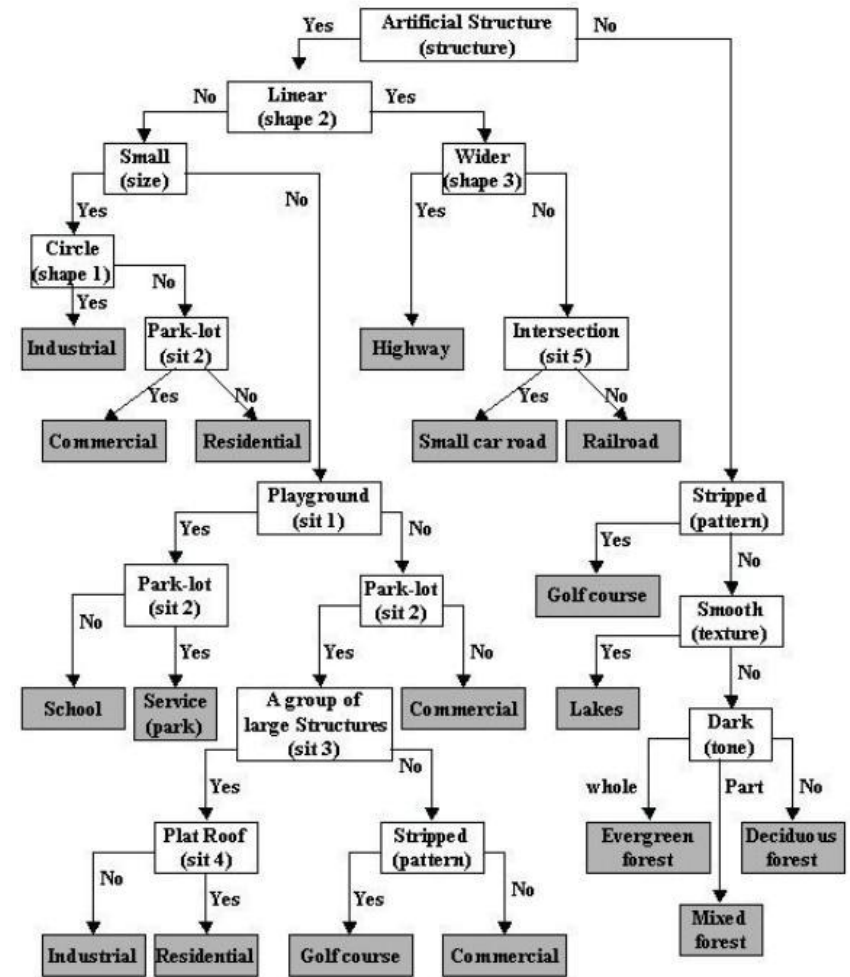
- Calculations can be complex

- Proper order of operations (e.g. canopy % vs transect depth)
    - Missing datapoints handled appropriately
    - Are *all* possible scenarios covered?
    - Rounding?

- QA can't build the calculations; prevents repeated mistakes

- Testing

- Test all scenarios, you'll still miss some





# Cautions, cont.

- Validation (QA)
  - Still must be done, even if it's just a formality
    - Verify accurate transfer to database
    - Verify calculations; may catch rare errors
    - There are still instances that are simply impractical/impossible to account for
- NELAC-if applicable



# Complexity

**Simple**

**Complex**

More

**Errors**

Less

Less

**Consistency**

More

More

**Time**

Less

Less

**Testing**

More

More

**Flexibility**

More

Less

**Maintenance**

Less



COMMITTED TO SAFE, CLEAN, ENJOYABLE CREEKS AND RIVERS

# Techniques

- Data Validation

- Restrict data:

- Specific Values
    - Range of Values
    - Dependent restrictions

- Prevent unreasonable/impossible value (e.g. longitude in Alaska)
- List, freeform, or list with freeform option

Location of Transect: Transect #		1			
Latitude		28.412340		Longitude -98.123450	
Left Bank Erosion	Instream	% Gravel or	Macrophyte	Algae	
Left Bank Slope (°)	Potential (%)	Cover (%)	Larger (>2mm)	Abundance	Abundance
105	105	40	85	Absent	Rare

Cover Types	
Undercut Banks	Yes
Gravel	Yes
Macrophytes	
Woody Debris	

Microsoft Excel

Percentages must be between 0 and 100.

[Was this information helpful?](#)

Dominant Substrate: (89844)		Stream
1-Clay; 2-Silt; 3-Sand;		1)
4-Gravel; 5-Cobble; 6-Boulder;		3)
7-Bedrock; 8-Other		
Reach Length of Stream	1	Habitat
Evaluated (89884)	2	1)
Channel Obstructions or	3	3)
Modifications	4	Restoration
Observed Stream Uses	5	1)
Depth of bottom at sample point	6	3)
(82903)	7	Land
Stream Segment Number	8	1)
	none	3)
	0.34	
	1011	





# Techniques, cont.

- Conditional Formatting

- Format field based on

- Value
    - Another value
    - Complex formula

- Flagging

- Often preferable to data validation to prevent blocking valid entries

Sample Number		XX#####	AB12345	AB123456	AB12347	
Station ID			12870	12871	12872	
Collection Date		MM/DD/YYYY	8/28/2018	8/28/2018	8/28/2018	
Collection Time		HH:MM or HHMM	9:45	10:26	11:31	
Sample Collector Initials (All)		XX/XX/XX/XX	DK/CV	DK/CV	DK/CV	
Field Dissolved Oxygen	00300	x.x	0.2	7.1	18.4	
Field Water Temperature	00010	xx.x	28.31	28.3	0.1	
Field pH	00400	x.x	12.1	7.8	8.12	
Field Chlorine Residual	50060	x.xx	Not Recorded	Not Recorded	Not Recorded	
Field Conductivity	00094	xxx or xxx0	541	1031	10.0	
Secchi Depth	00078	x.x or 0.xx	1.31	1.5	0.91	
Days Since Last Precipitation Event	72053	<1, 1-75, >75	12	0	<1	
Flow Severity	01351	1=No Flow, 2=Low, 3=Normal, 4=Flood, 5=High, 6=Dry				

	1	2	3	Total (mean weighted by kick distance)		
<b>Benthic Kick Distance</b>	2	3	2	7	m	
<b>Gravel Substrate</b>	80	90	80	84.3	%	Due to rounding, total % may not = 100
<b>Sand Substrate</b>	15	10	10	11.4	%	
<b>Soft Bottom</b>	10		10	5.7	%	
<b>Macrophyte bed</b>	0		0	0.0	%	
<b>Snags and brush</b>	0		0	0.0	%	
<b>Bedrock</b>	0		0	0.0	%	



# Where SARA is

- In-House Development (Excel)
- Rugged Field Tablet with full Windows capabilities (tablet PC)
- Routine Field/Flow
  - Single COC and Electronic file (up to 9 samples)
- Biological Sampling
  - Single COC for Nekton, Habitat, Benthics
  - Electronic file for each (3 total)
  - Complex calculations and error-checking



# Where SARA is

- All related forms maintained as controlled documents
  - Track what is changed and why; prevent accidental changes or loss of original
- Updates
  - Yearly updates (if needed) provide fixes for low-priority bugs, refinement of features, and new features (becoming less common)
  - Immediate updates for major functionality or calculation errors (rare)





# Where SARA is

- Process

- Prep

- Fill out and print paper COC
    - Download electronic forms to tablet (ensure someone who will be present is logged on!)
    - Pre-fill applicable fields in electronic forms (e.g., date, collectors, site #)

- Collection

- Collect data; record in electronic form and/or COC as applicable
    - Relinquish COC to appropriate staff

- Post-processing

- Review electronic forms for completeness
    - Create CSV
    - Upload to database (LIMS using Labworks 6.8)
    - Store electronic forms in designated server location



# SARA's Forms

## Routine Field COC

- Paper COC covering up to 9 samples

## Routine Field Data

- Limited macros (shortcuts, CSV creation)
- Heavy conditional formatting for flagging
  - Empty fields
  - Improper precision or format
  - Out of range
- Pre-fill basic info; review, store, and upload on return



# SARA's Forms

## Biological COC

- Simple form
- Single biological sample set; includes separate line for each sample group
  - Electrofishing
  - Seining
  - Nekton Metadata/Scoring
  - Benthic Macroinvertebrates
  - Physical Habitat



# SARA's Forms

## Nekton Community

- Captures electrofishing, seining, and metadata/scoring in one file
- Designed to roughly mimic “tick mark” style counting
- Most entry via touch screen
- Count recording heavily macro-based
- Live scoring
- Pre-fill known data (drainage size, stream order, site info, date)
- Minimal post-collection entries



# SARA's Forms

## Benthic Macroinvertebrate Community

- Basically a calculation sheet
- Limited macros
- Comprehensive reference list (STORET, FFG, etc.)
- Enter STORET code and count; all taxa data and calculations automated
- RBI Scoring Automatic
- Queries LIMS to ensure most current list of reported taxa for upload
- Minimal data collected aside from ID and enumeration





# SARA's Forms

## Physical Habitat

- Limited macros; for extreme data validation and shortcuts
- Conditional formatting for flagging
- Data validation – drop down lists and impossible values
- Complex calculations for missing/nonnumerical entries
- Pre-fill known, static values
- HBI scoring manual



# What Next?

- Intensive Nekton Survey (similar to TIFP sampling)
  - Based on original nekton form + measurement recording
  - Master File
  - Data Collection File
- Electronic COC?
  - Must capture NELAC acceptable signature





# Questions???

**Doug Knabe**  
**San Antonio River Authority**  
**[dknabe@sara-tx.org](mailto:dknabe@sara-tx.org)**



COMMITTED TO SAFE, CLEAN, ENJOYABLE CREEKS AND RIVERS

# Final Tips for Developers

- If your macros refer to a specific cell or range:
  - Create a “middleman” page with a list of the cell references (and a description), then refer to THAT cell to obtain the range within the macro; this simplifies updates if a cell reference changes
  - Can populate the “middleman” using formulas to provide auto-update if a cell is moved
- Excel does not use scientific rounding, but:
  - Rounding in a macro DOES use scientific rounding (i.e. create custom function, must enable macros)
  - But, in a **formula** (no macros required),
    - **V** = value, **P** = # of decimal places
    - **IF( TRUNC(V,P) = (V - 0.5\*0.1^P), MROUND(V,2\*0.1^P), ROUND(V,P))**



# Development timeline

- Routine:
  - Quick initial development
  - Parallel testing for 2 months
  - Bug fixing as needed (rare)
  - Recent developments (full error checking) ~4 weeks as time allowed
- Biological:
  - One “off-season” (Sept-March) as time allowed
  - 2 events for parallel testing and bug fixing
  - Continued refinement and bug fixing as needed to streamline processes and further reduce user-error (yearly, if needed)





# Tablet Information

- xTablet T1200; XT125-2010
- Temperature: -20 to 60°C
- Humidity: 5 to 95%
- IP65 (dust tight, water resistant)
- Bluetooth capable
- ~5 lbs
- 11.5 hr battery life

